**Chapter 11 exercises**

***This homework is being posted late in the week. Rather than extend the due date, which will interfere with the next homework, I am including just a couple of questions which you should be able to answer quickly.***

*Submit by sharing your copy (renamed to your name) with* [*denys.katerenchuk@gmail.com*](mailto:denys.katerenchuk@gmail.com) *and* [***rebeccalevitan@share.brooklyn.edu***](mailto:rebeccalevitan@share.brooklyn.edu)

|  |  |
| --- | --- |
| **Problem** | **Technique** |
| Vanishing gradients | Batch Normalization |
| Dying ReLU | Leaky ReLu |

What is Batch Normalization? What is it good for? What is its downside?

Batch Normalization standardizes and normalizes input data [in batches]. If you put a batch normalization layer for every hidden layer, you remove the problem of vanishing (and exploding) gradients and prevent it from coming back later in the training process.

BN also reduces the amount that suboptimal weight initialization hurts the neural net, and acts as a regularizer overall.

It does increase the prediction time required because of the added calculations, though is very widely used in practice regardless.